REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
05-05-2014	Final	15 Feb 2011- 14 Sep 2013

4. TITLE AND SUBTITLE Multi-Robot Teaming - MAGIC 2010 Second Place - University of Pennsylvania	**5a. CONTRACT NUMBER** FA23861114026
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER **61102F**
6. AUTHOR(S) Dr. Daniel D. Lee	**5d. PROJECT NUMBER**
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Pennsylvania 200 S. 33rd Street Philadelphia, PA 19104 United States	**8. PERFORMING ORGANIZATION REPORT NUMBER** N/A
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) AOARD UNIT 45002 APO AP 96338-5002	**10. SPONSOR/MONITOR'S ACRONYM(S)** AFRL/AFOSR/IOA (AOARD)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S) AOARD-114026

12. DISTRIBUTION/AVAILABILITY STATEMENT

Distribution A: Approved for public release, distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

In this report, we describe the technical approach and algorithms that have been used by the University of Pennsylvania for multi-robot perception, planning, and control in the context of the MAGIC competition and in ensuing years. We have constructed and deployed a multi-vehicle robot team, consisting of intelligent sensor and disrupter UGVs that can survey, map, recognize, and respond to threats in a dynamic urban environment with minimal human guidance. The custom hardware systems consist of robust and complementary sensors, integrated electronics, computation, and highly capable propulsion and actuation. The mapping, navigation, and planning software is organized hierarchically, allowing autonomous decisions to be made by the robots while enabling human operators to interact with the robot team in an efficient and strategic manner. The ground control station integrates information coming from the robots as well as metadata feeds to focus the attention of the operators and respond rapidly to emerging threats.

15. SUBJECT TERMS

Multi-Agent Technology, Unmanned Vehicles, Robotics

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Brian Sells, Lt Col, USAF
U	U	U	SAR	34	**19b. TELEPHONE NUMBER** *(Include area code)* +81-3-5410-4409

Distribution A: Approved for public release, distribution is unlimited.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
05 MAY 2014	**Final**	**15-02-2011 to 14-09-2013**

4. TITLE AND SUBTITLE
Multi-Robot Teaming - MAGIC 2010 Second Place - University of Pennsylvania

5a. CONTRACT NUMBER
FA23861114026

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
61102F

6. AUTHOR(S)
Daniel D. Lee

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Pennsylvania,200 S. ,33rd Street,Philadelphia,PA,19104

8. PERFORMING ORGANIZATION REPORT NUMBER
N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AOARD, UNIT 45002, APO, AP, 96338-5002

10. SPONSOR/MONITOR’S ACRONYM(S)
AOARD

11. SPONSOR/MONITOR’S REPORT NUMBER(S)
AOARD-114026

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
In this report, we describe the technical approach and algorithms that have been used by the University of Pennsylvania for multi-robot perception, planning, and control in the context of the MAGIC competition and in ensuing years. We have constructed and deployed a multi-vehicle robot team, consisting of intelligent sensor and disrupter UGVs that can survey, map, recognize, and respond to threats in a dynamic urban environment with minimal human guidance. The custom hardware systems consist of robust and complementary sensors, integrated electronics, computation, and highly capable propulsion and actuation. The mapping, navigation, and planning software is organized hierarchically, allowing autonomous decisions to be made by the robots while enabling human operators to interact with the robot team in an efficient and strategic manner. The ground control station integrates information coming from the robots as well as metadata feeds to focus the attention of the operators and respond rapidly to emerging threats.

15. SUBJECT TERMS
Multi-Agent Technology, Unmanned Vehicles, Robotics

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**Same as Report (SAR)**	**34**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

UPenn Multi-Robot Unmanned Vehicle System (MAGIC)

AFOSR Final Report

PI: Daniel D. Lee*

GRASP Laboratory

University of Pennsylvania

Philadelphia, PA 19104

Abstract

In this report, we describe the technical approach and algorithms that have been used by the University of Pennsylvania for multi-robot perception, planning, and control in the context of the MAGIC competition and in ensuing years. We have constructed and deployed a multi-vehicle robot team, consisting of intelligent sensor and disrupter UGVs that can survey, map, recognize, and respond to threats in a dynamic urban environment with minimal human guidance. The custom hardware systems consist of robust and complementary sensors, integrated electronics, computation, and highly capable propulsion and actuation. The mapping, navigation, and planning software is organized hierarchically, allowing autonomous decisions to be made by the robots while enabling human operators to interact with the robot team in an efficient and strategic manner. The ground control station integrates information coming from the robots as well as metadata feeds to focus the attention of the operators and respond rapidly to emerging threats.

1 Introduction

The goal of the 2010 Multi Autonomous Ground-robotic International Challenge (MAGIC 2010) was to bring together robotics experts from all over the world to participate in a unique competition. Jointly organized by the Defence Science and Technology Organisation (DSTO) in Australia and the Research Development and Engineering Command (RDECOM) in the USA, the challenge called for researchers to develop and successfully field a team of robots which could explore and map a large dynamic urban environment, as well as locate, classify and respond to threats. The purpose of the challenge was to significantly accelerate the development of autonomous and unmanned robotic teams, which could operate effectively with limited guidance from human operators. In the following sections, we report on the technical approaches of the UPenn team, consisting of students and faculty from the School of Engineering and Applied Science at the University of Pennsylvania.

1.1 Problem statement

The objectives involved designing and constructing a team of ground robots that could operate in unknown environments and execute complex missions with minimal human interaction. It was set up to challenge the individual robots as well as to test their ability to work in teams and execute tasks explicitly requiring simultaneous participation of multiple autonomous agents. Mobility was not the primary emphasis of the challenge, however reasonable capabilities to traverse uneven indoor and outdoor terrains were expected. Each robot had to demonstrate the ability to:

- navigate rough terrain, grass, sand, bumps, and clearing 10 cm curbs
- sense and explore the surrounding environment using cameras, laser scanners, and other approved active and passive sensors
- integrate sensor data to construct an accurate metric representation of the environment
- optimize trajectories and safely navigate in the environment
- communicate its collected information to other robots and the ground station

- accept human control inputs and override commands transmitted wirelessly
- detect objects of interest (OOIs) and report them to the ground station
- perform a neutralization procedure upon confirmation from a human operator

(a) indoor barn environment (b) maze environment (c) static OOI

Figure 1: Representative images from the environment. The static objects of interest (OOIs) were red barrels with a predetermined surrounding danger zone. Robots that entered the danger zone before neutralization were disabled.

(a) civilian (b) mobile OOI

Figure 2: A civilian and a mobile OOI. The mobile OOIs had to be neutralized, but not in the proximity of civilians. Mobile OOIs were neutralized by two robots simultaneously tracking the OOI for 30 seconds at a specified distance.

In addition to demonstrating performance of individual robots, the teams also had to show how they could manage and coordinate behaviors across the different robots. The complex missions often called for certain robots to switch modes of operation, from full autonomy to obeying high level human commands. These modes included exploration, waypoint navigation, threat detection, maintaining distance from static and moving targets, OOI neutralization, and failure recovery modes. Due to the limit on the number of operators (maximum of two human operators per team)

and the large number of robots (up to nine robots for the UPenn team), the system had to operate at a semi-autonomous level and seamlessly allow operator monitoring and access to individual or groups of robots. Additionally, certain tasks implicitly or explicitly required multiple robots working together. For example :

- exploring large unknown environments and generating globally consistent and GPS-registered metric maps
- executing complex threat neutralization procedures requiring simultaneous involvement of multiple robots
- assigning and tasking robot groups to handle newly discovered events and threats

The team aspect of the problem was the most important, since projects and challenges involving single robots have been held numerious times in the past: DARPA Urban Challenge (Bohren et al., 2008), LAGR (Vernaza et al., 2008), among others. Novel approaches to efficiently interact with a team of semi-autonomous robots without requiring a large number of well-trained human operators needed to be developed. The results of this research could be used to minimize the risk to human lives in dangerous search and surveillance missions, such as in large-scale disaster search and rescue scenarios.

2 Overall System Architecture

We formulated a broad solution that encompasses a wide range of spatial and temporal scales. Our solution uses a carefully constructed *hierarchical* decomposition of perceptual, planning, and control algorithms shown in Figure 3. This hierarchical solution allows for efficient high-level interaction from the human operators while simultaneously allowing the robots to operate in an autonomous manner.

Our solution uses a bottom-up hierarchical organization of sensing and mapping tasks, integrating low-level sensor readings at fast time scales from individual sensor robots into a global overhead view for the human operators. This high-level display of the world is then overlaid with metadata

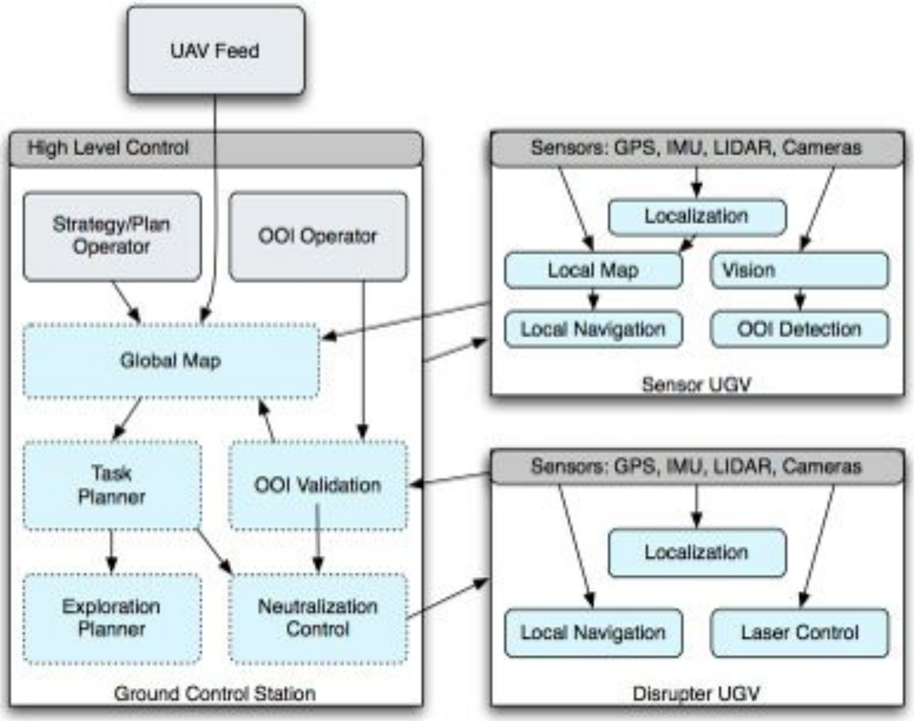


Figure 3: Our approach uses a hierarchical decomposition of perceptual, planning, and control tasks with high-level human operator commands integrated with low-level autonomous robot algorithms.

feeds from simulated UAVs, as well as validated object-of-interest (OOI) positions. A top-down hierarchical planning and strategic control decomposition allows the human operators to efficiently issue high-level task commands that get propagated down a chain of planners and controllers to ultimately navigate the individual robots to their desired locations and execute appropriate actions.

Our overall system involves successfully integrating the various components to handle a complete mission such as the layout shown in Figure 4. Our system is comprised of the following subsystems:

- Sensor UGV: Mobile UGVs with LIDAR and camera sensors, GPS, and IMU perform low level localization, mapping, and visual detection tasks. Each sensor UGV is capable of autonomous navigation to explore unknown terrain, as well as tracking mobile OOIs.

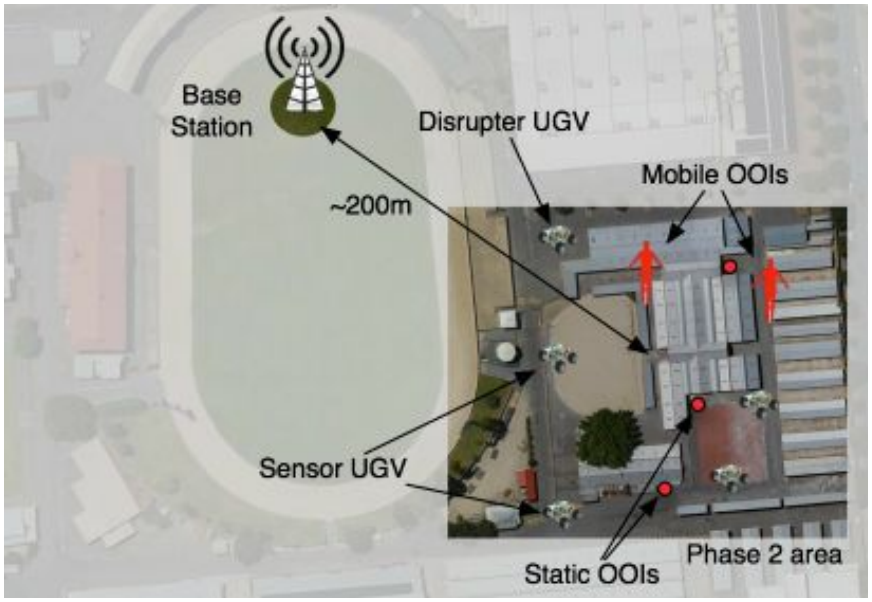


Figure 4: Overview of the mission scenario in the Adelaide Fairgrounds, showing locations of the sensor and disrupter UGVs, along with the known positions of static and mobile OOIs in the environment.

- Disrupter UGV: Highly mobile UGVs with sensors and a pan-tilt laser pointer used to neutralize static objects of interest (OOI). These UGVs require human operator approval to verify potential OOIs.

- Wireless communications system: A long-range wireless communications system over unlicensed bands provides redundant communication links between the UGVs and the ground control station. The system utilizes some of the sensor UGVs as relay repeaters to facilitate communications in difficult RF environments.

- Ground control computers: A cluster of multi-processor computers integrate all incoming sensor information from the UGVs as well as the simulated UAV feed into a global map with static and mobile OOI locations. This map is used by the human operators for strategic planning and for display to the challenge judges.

- Strategy/Plan control station: A high-resolution graphical interface is provided to allow the Strategy/Plan operator to input areas of interest for exploration and mapping, and to allocate tasks to different robots.

- OOI validation control station: Potential OOIs identified by UGVs are displayed in both omnidirectional and high-resolution images for validation by the OOI operator.

- Simulated UAV feed: Provided by the competition organizers, this feed gives additional information for situational awareness, such as potential locations of OOIs, but this information was inaccurate as well as incomplete.

The modules in each of these subsystems are organized hierarchically, and are connected to each other using fault-tolerant interprocess communications protocols. Common to the UGV platforms are modules that integrate sensory information at a local level for pose estimation and automatic object recognition. Sensor UGVs are specialized to build local maps of obstacles and static OOIs and to accurately track mobile OOIs. This information is then forwarded to the ground control stations where human operators can make high-level command decisions to neutralize OOIs or avoid certain regions. Disrupter UGVs can then be tasked to navigate to validated OOIs to initiate the neutralization process.

The ground control station incorporates multi-core processors that integrate the various lower-level data streams. A hierarchical mapping module provides a global map view of the environment by aggregating and registering the local maps from the individual robots. An overlay with UAV feed information is then used to generate real-time updates of static object locations as well as dynamic object movements. The high-level maps generated by these modules are available for display and monitoring for the human operators.

This information is utilized by the operators so that they can issue high-level commands to a series of planning algorithms. An intuitive point and click interface modeled after real-time strategy computer games allows for rapid human input to control robot task allocation as well as strategic planning. A series of exploration and task assignment planners then compute the optimal routes and actions for the robot team by balancing exploration, mapping, and threat neutralization objectives. Low-level commands are then relayed by the system back to the individual UGVs for execution.

3 Ground Vehicle Components and Systems

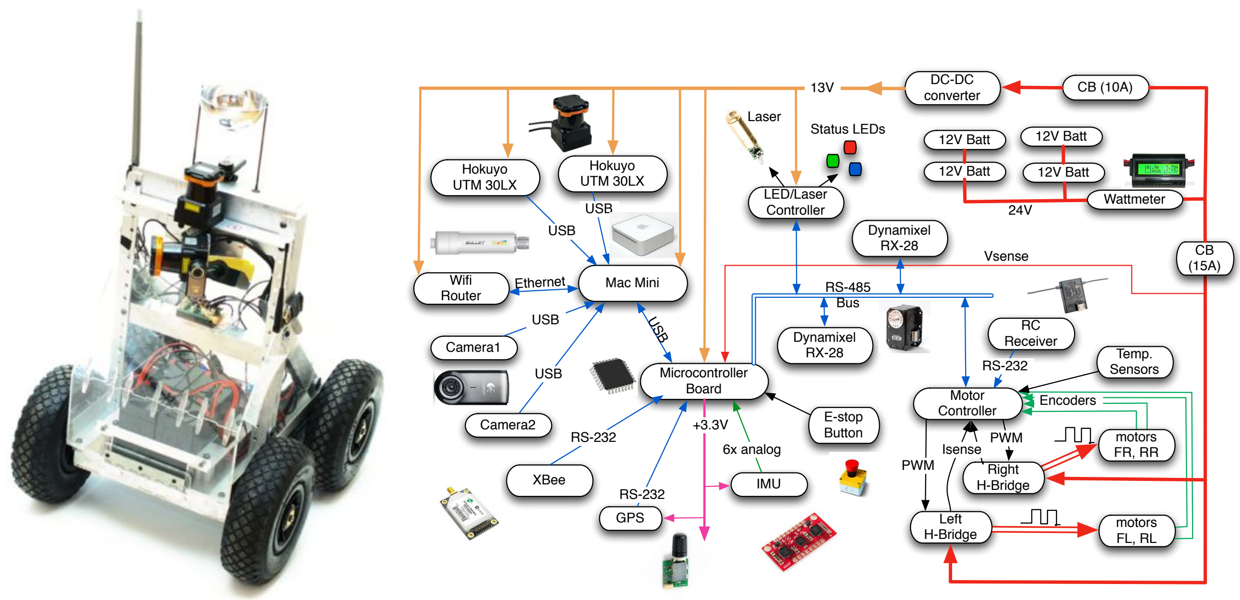


Figure 5: UGVs use an all-terrain aluminum robot base with complementary sets of LIDAR and camera sensors, GPS, IMU, battery-powered embedded computers and wireless connectivity.

Our ground robots are built on a lightweight, all-terrain robot vehicle base as shown in Figure 5. The vehicle base is constructed from welded aircraft-grade aluminum, with a high current DC motor drive system connected to a set of four rugged 25 cm wheels, capable of traversing over 10 cm tall obstacles and speeds up to 2 m/s on flat terrain.

The power system uses two sets of rechargeable 2×12V batteries that are configured so that they may be hotswapped in the field, eliminating the need to reboot the robot. The power is distributed via a series of switching DC-DC converters and customized control electronics to a large set of complementary sensors. The sensor suite for each robot consists of the following:

- Horizontal scanning Hokuyo LIDAR detector returning laser ranges from up to 30 m away
- Vertical scanning Hokuyo LIDAR detector on a panning servo motor returning ground laser ranges
- Omnidirectional catoptric color camera (Logitech C905 and hemispherical mirror)

- Panning frontal view color camera (Logitech C905)
- Hall-effect motor encoders for proprioception
- Custom built 6 degree of freedom strapdown inertial measurement unit integrating MEMS-based gyroscopes and accelerometers. An Atmel microcontroller is used to perform sensor fusion using attitude estimation as described in (Trimpe and DAndrea, 2010).
- 50 channel helical GPS receiver

The disrupter UGVs have an additional pan-tilt degree of freedom that allows for aiming an eye-safe laser pointer to neutralize identified static OOIs. Computational power onboard the robots is provided by an embedded dual-core Mac Mini (2.66 GHz, 4GB RAM) computer running Linux on a solid-state drive, with USB connections to the various sensors and microcontrollers. A 802.11g WiFi interface enables high bandwidth network connectivity to the ground control computers, while an Xbee radio link provides a redundant communications channel for emergency control and safety purposes.

4 UGV Autonomy and Coordination Strategy

We constructed and fielded nine UGVs (six sensor UGVs and three disrupter UGVs) as shown in Figure 7. In order for only two human operators to effectively control this team of robots, there must be a great deal of autonomy built into the system to reduce the level of cognitive overload on the human operators (Hsieh et al., 2007; Michael et al., 2008). In this section, we summarize the high-level functionality of our system and show how it enabled the two human operators to control the team of robots.

4.1 Robot functions and operator access levels

Each robot in the group was equipped to sense, map, and autonomously navigate the surrounding environment in an independent and safe manner. Our high-level interface software provided control access for the operators to interact with the individual robots or a group. The high-level commands

include:

- Go-to-point: select a particular robot and assign a goal position. The global planner generates a rough trajectory, which is sent to the robot and whose on-board local planner refines the trajectory for smooth navigation. Once the smooth trajectory is generated, the robot navigates towards the goal and re-plans accordingly as new obstacles are discovered. See Section 6 for more details on the motion planner.
- Explore: uses a high-level planner to assign one or more robots to explore a specified area or an area containing a combination of "explore" and "avoid" regions. The exploration planner, running on the base station, generates goals for each robot and those are sent as waypoints for execution. During exploration, the OOI detection algorithm constantly processes the video from each robot, searching for potential OOIs.
- Track a moving OOI: operator clicks anywhere in the camera view of the robot and the selected object in the view becomes the target for tracking. As the target moves, the robot turns and adjusts the position of the servo, keeping the target in the field of view of the front camera
- Neutralize a static OOI: disrupter UGVs were equipped with laser pointers, which could be used to neutralize the static OOIs (red barrels). Before neutralization, the robot must confirm the operation with the human operator.

4.2 Multi-robot exploration

One of the major tasks in the challenge was exploring large unknown regions. For example, consider the two-robot test mission shown in Figure 8. Here, two robots are initialized next to each other and are tasked to explore all accessable space of this indoor environment. In this scenario containing over 100 m of corridors, the two robots complete the exploration task in under four minutes.

Single robot exploration problem has been previously studied and most successful solutions are based upon maximizing the expected information gain. New information during exploration repre-

sents sensor data that reveals previously unknown regions of the map. In our case, the exploration map is a 2D grid with 20 cm resolution, with each cell in the grid containing a binary value: known or unknown. The goal of the exploration planner is to predict where the largest information gain is located and to send the robots towards these frontiers (Yamauchi, 1997).

Figure 9 shows the frontiers for a single and combined multiple robot exploration scenario. Here a large outdoor area has been partially mapped, and no static OOIs have been located on the ground. The planner automatically determines the most promising frontier areas to explore, and allocates locations for each individual UGV to navigate. This is done by optimizing the overall information gain in traversing states j

$$IG(j) = \sum_{k \in vis(j)} q_k \tag{1}$$

where q_k is inversely related to the amount of knowledge about state k that is reachable from state j, as determined by the shortest paths computed from the current known map. These locations are then sent to a lower-level navigation planner onboard each robot to autonomously execute in real time.

Algorithms have been developed that extend frontier exploration scheme to robot teams (Visser and Slamet, 2008; Simmons et al., 2000; Zlot et al., 2002). In this challenge, however, it was critical to seamlessly integrate human operator input at various levels, including exploration. We introduced a centralized algorithm for multi-robot exploration that not only uses the frontier-based approach, but also incorporates soft and hard constraints on the individual robot positions that can be interactively set by human operators. In addition, the algorithm can handle constraints defining maximum and minimum inter-robot distances and heading bias (Butzke and Likhachev, 2011).

Additional constraints make the exploration problem more difficult to solve, so there was a need to make computing the solution more efficient. We needed to allow the human operators to quickly put in hints for the exploration planner to speed up performance in complicated scenarios. Figure 10 shows an example of how our team used the graphical user interface to help guide multi-robot exploration missions. The exploration map and these constraints are translated into a utility functions, which along with the cost map are used in the optimization to find the best goal states.

Section 7 describes the results of our exploration module with five robots exploring and mapping a large indoor environment for the Old Ram Shed Challenge.

4.3 Handling OOIs

During exploration, the robots may locate potential OOIs and alert the operators. Once a putative OOI has been located, a popup message is displayed to the OOI operator asking to verify and validate the identity of the OOI. If the human operator ignores this message, the sensor UGVs will automatically continue with their search tasks. However, if the identity of the static OOI is confirmed, the location of the OOI is forwarded to the central strategic control computer for display on the global map. At this point, the Strategy/Plan operator may choose to initiate a neutralization with a single click on the OOI location. The system automatically then cross-cues the nearest disrupter UGV to navigate towards the OOI location, simultaneously placing an avoidance region 3 meters around the OOI.

Once the disrupter UGV has reached a location at a safe distance away from the OOI neutralization zone, the OOI operator confirms visual acquisition of the correct target using images streamed from the disrupter UGVs forward camera (Figure 11). The sensor UGV cameras can also be used to increase situational awareness during the neutralization process. Once the OOI target has been centered, a command is relayed from the ground control station to the disrupter UGV to start neutralization by activating its laser. When the neutralization process is complete, the OOI is marked as a neutralized OOI in the global map, and the UGVs automatically return to their previous tasks.

On the other hand, when a dynamic OOI is discovered by a UGV or indicated in the UAV feed, two sensor UGVs are tasked to navigate to locations surrounding the path of the dynamic OOI. Once the UGVs have reached their desired locations, a visual confirmation is given by the OOI operator indicating the identity and location of the dynamic OOI. At this point, a command is relayed to the sensor UGVs to switch into an autonomous tracking mode. In this mode, the robot cameras are servoed to follow the moving OOI. Once visual lock has been achieved, the OOI operator checks

the surrounding area for non-combatants and begins neutralization if the area is clear (shown in Figure 12). When neutralization has finished, the sensor OOIs are switched back into their search and exploration mode.

5 Sensors, Processing and Mapping for UGVs

5.1 Sensor processing

The onboard perceptual system processes information from the multiple cameras and LIDAR sensors on each UGV. There are two USB cameras mounted on each robot to provide a complementary set of visual images. A hemispherical mirror (taken from an inexpensive plastic holiday ornament) is used with one of the cameras to provide a 360 degree field of view enabling rapid search for potential OOIs and regions of interest around the robot (Geyer and Daniilidis, 2003). An example of an unwarped omnidirectional image taken by a UGV is shown in Figure 13. A second front facing camera is mounted on a panning servo that allows for closer inspection of any identified areas of interest in the omnidirectional view.

To assist in the labeling of static and mobile OOIs, each robot performs color segmentation and object detection routine on the omnidirectional and front facing camera images. This algorithm analyzes connected regions of high Cr values in the YCbCr image space, and returns bounding boxes of potential regions of interest along with corresponding scores for each region. These regions are presented to the OOI human operator superimposed on the images, with each region's bounding box highlighted according to the saliency of the region's rank. Depths of potential objects are computed by correlating region size with information from the UGV LIDAR sensors (Figure 11).

We have implemented a calibration procedure to register the camera frame to the LIDAR and servo frames, which is required in order to achieve accurate 3D mapping. The appropriate procedure is described in (Naroditsky et al., 2011) and involves calibrating the reference frames from multiple observations of a known calibration pattern. Scanning LIDAR depth and intensity data can be used to align both geometric and color features. Figure 14 shows how a well-calibrated system can be

used generate consistent colored 3D maps, similar in performance to the RGBD sensors (Henry et al., 2012), but with the ability to work in sunlight as well. This procedure was not enabled during the competition, since accurate 3D color point clouds was not required for scoring. However, a similar procedure was used to align the sensor frames for approximate depth extraction on the visual features.

5.2 Localization and mapping

The main localization and mapping modules rely on the Hokuyo UTM30-LX LIDAR sensors, configured as shown in Figure 6. Each robot is equipped with two LIDARs: one is fixed in a horizontal orientation, and the other scans vertically with the ability to rotate the scan plane over 120 degrees with a small servomotor. The horizontal sensor is fixed with respect to the robot frame, and its long-range returns are used for 2D localization and occupancy grid mapping in conjuction with a differential drive motion model. The vertical sensor gathers dense information about the ground in front of the robot and helps determine the traversability of the surrounding terrain. As the sensor continuously scans the frontal area, dangerous three-dimensional obstacles such as low-lying structures and curbs can be robustly detected.

We have explicitly chosen this configuration over the "nodding" motion used by other researchers, since it allows the perceptual system to get faster map updates in critical areas in front of the robot (immediately in front of the wheels to five meters ahead of the robot). Vertical scans were not used for pose estimation, due to the uncertain geometry of many prevalent features such as bushes, trees, and dynamic OOIs. The horizontal LIDAR provided accurate local pose tracking via scan matching at 40 Hz and the vertical scans were used to accurately populate the obstacle map.

The range data acquired by these two sensors is integrated with measurements provided by the motor encoders and onboard inertial measurement unit. In this manner, the three-dimensional orientation of the robot can be tracked with very low latency. The odometry and inertial readings are combined with laser scan matching to localize the robot using a probabilistic Rao-Blackwellized particle filter to properly keep track of changes in orientation as well as translational motion (Ver-

naza and Lee, 2006). This filter relies upon a factorized representation of the underlying pose:

$$P(x, y, \theta) = \sum_i \delta(\theta - \theta_i) N(x, y|\theta_i) \tag{2}$$

where the heading angle θ is sampled and the translational degrees of freedom (x, y) are conditionally independent given the heading.

Simultaneous localization and mapping (SLAM) techniques, based on highly optimized scan matching routines similar to techniques discussed in (Olson, 2008), are employed to build a local map of the environment surrounding the robot. Readings from the two complementary LIDAR sensors are used to update a probabilistic grid map where both traversable ground and obstacles are recorded. An example of such a map generated by a single and multiple UGVs is shown in Figure 15. In this figure, traversable ground as determined by the vertical LIDAR is shown in light colors, and obstacles such as walls and low-lying objects are shown in yellow, red and black.

Each robot keeps track of its local map and uses it for navigation. Map updates from individual robots are sent to the base station and fused based on the consistency of the maps and GPS readings, as discussed in Section 7.

6 Motion Planning

The Ground Control Station (GCS) provides the robot with a rough 8-connected path at 20 cm resolution, either from the exploration planner or from the “Go To Point” 2D Dijkstra search. This path may be difficult to follow due to potential sharp turns that are inherent to these types of planners. Our system produces a smoother path that is easier to follow while maintaining proximity to the original one (in exploration, the generated paths often provide significant information gain, so following accuracy is important for consistency of the system). Our local planner uses an A*-based method on a (x, y, θ) lattice graph to produce smooth paths which are more easily executed (Likhachev and Ferguson, 2008). An example graph is shown in Figure 16.

Normally on an 8-connected grid, each state is connected to its 8 immediate neighbors. On the other hand, in a lattice graph the successors of a state are the end points of different “motion

primitives". A motion primitive is a short dynamically feasible motion, derived from the dynamic differential-drive vehicle motion model. The set of motion primitives are customized for the UGV to be easy to follow with a local planner, resulting in smooth paths that take into account the robot's motion constraints.

To encourage the planner to stay close to the original GCS provided trajectory, we create a special costmap. The costmap we use is comprised of two components. The first component keeps the robot away from obstacles by linearly increasing the cost of cells as the distance computed using the occupancy map to the nearest obstacle decreases. This component will discourage the robot from driving near obstacles but the planner is still able to get through a tight area if needed. The second component of the costmap encourages the planner to stay close to the original path given by the GCS. This is done by increasing the cost of cells as they deviate from the GCS path above a set distance threshold. The final costmap is determined by taking the maximum of each cell in these two maps.

After creating the costmap, we find a solution using an anytime search-based algorithm, ARA* (Likhachev et al., 2003). This planner finds a suboptimal solution quickly and then improves the solution as time allows. Given enough time the planner will find the optimal solution with respect to the costmap and motion primitives, but at any time after an initial fast search, it can return a solution with bounded sub-optimality. This means that the robot can start executing the trajectory even before the optimal solution is found with guarantees that it will not be too far off from the optimal trajectory.

7 Operations in GPS-denied Environments

Each UGV is equipped with a relatively low-cost GPS receiver (50 channel D2523T module) which provides absolute position information at 1 Hz. However, because GPS is not available indoors and can be highly unreliable due to occlusions and multi-path effects, we designed most of our systems to operate without relying upon these measurements.

UGV ID	Distance	Time	Average Speed
1	448 m	34.6 min	0.216 m/s
2	388 m	26.4 min (power failure)	0.213 m/s
3	375 m	34.6 min	0.181 m/s
4	337 m	31.9 min	0.176 m/s
5	368 m	34.6 min	0.177 m/s

Table 1: UGV travel distances, times, and average speeds during the Old Ram Shed Challenge.

Instead, our operations rely upon a globally consistent map built using a hierarchical map registration algorithm that does not use GPS. First, local maps are built on each UGV using only odometry, IMU, and LIDAR readings. These local maps are then sent incrementally to a higher level map registration module which merges the incoming local maps from each robot into a globally consistent map by maximizing the likelihood of the map cell occupancy statistics, using methods similar to (Reid and Bräunl, 2011).

Figure 18 shows an example of the globally consistent map computed by merging several UGV maps together in the Old Ram Shed Challenge where GPS was not available. This map was simultaneously constructed using five UGV robots, which traversed the environment via trajectories computed by the exploration module and local planners. The trajectory statistics of each robot in constructing the map is given in Table 1.

GPS measurements are incorporated only after the lower level mapping modules update a globally consistent map. This final stage of processing registers the map to a global reference frame using the very sporadic GPS readings when there are many satellites in view and the horizontal dilution of precision is very accurate. This is accomplished by finding the optimal rigid transformation T^* between the global map's coordinate frame and absolute UTM coordinates by optimizing the squared error between the good GPS readings and map coordinates:

$$T^* = \arg\min_{T} \sum_{i} |\vec{x}_{GPS} - T\vec{x}_{map}|^2 \tag{3}$$

Thus, GPS readings are only used to determine a rigid coordinate transformation between each robot's local frame and absolute UTM coordinates. Unreliable GPS readings will only skew this

coordinate transformation, and will not affect the overall quality of the map used by the human operators for mission purposes.

8 Human Robot Interface

8.1 Metadata fusion

A priori information about the environment is merged with the global map built using the UGV sensors in order to provide contextual guidance to the human operators. This process is illustrated in Figure 19 where Google Earth imagery is shown overlaid on an indoor/outdoor global map constructed by the UGVs. First, image processing algorithms are applied to the overhead images to extract edge and texture information. The locations of these features are then registered with the current occupancy map, and the probabilities within the corresponding cells in the map are updated accordingly. These maps can then be used for preplanning mission operations by highlighting areas of interest (Ferguson et al., 2008).

An additional ground control software module is responsible for connecting to the simulated UAV data feed via TCP network protocols, and converting transmitted OOI information to the global map module. These transmissions are logged and registered for display on the map, just as the data coming from the individual UGVs are recorded. Thus, in our framework, incorporating metadata from a UAV is no more complex than fusing information from the different sensor UGVs. An overlay of the UAV feed can be switched on and off by the Strategy/Plan operator. This information can then be used to quickly determine desirable exploration or undesirable avoidance regions in the global map.

8.2 Situational awareness

Fast and faithful situational awareness is vital in dynamic environments where control decisions are time critical and carry immediate consequences. In our system, relevant percepts such as OOI locations are passed to the navigation modules, allowing each UGV to avoid hazardous situations

without human control. In an ideal perception system, the robot would automatically detect, recognize, and track all lethal OOIs in real time with high accuracy. Due to the challenges posed by the competition however, it was necessary to weigh the trade offs between human interaction, speed, and accuracy. While object detection and recognition are well-studied problems in computer vision (Edelman, 1997; Viola and Jones, 2001), their performance is compromised by real-world effects such as color constancy and occlusion. These issues can be quite severe in practice as the competition environment was cluttered with highly variable and extreme lighting. Therefore, we focused on designing a user interface that made available as much information as possible while guiding the operator's limited attentional resource (Vanrullen et al., 2004).

The OOI operator's attention is primarily focused on the center region of the main screen display. Patches of enhanced resolution (Fig. 20c) are presented to the user in order of saliency as a red object. Specifically, pixel patches that have the closest distance to the OOI color and are within a reasonable size range are selected and presented. Mouse clicks or key presses for the corresponding patch can be used to quickly focus, initialize tracking, or confirm the identity as an OOI in the detected region of interest, which is then presented in primary focus.

Two UGV's can be selected at a time for primary focus, with an omnidirectional view containing clock direction markings and a front panning view displayed in the top region of the screen (Fig. 20a). Clicking on any omni view (Fig. 20b) immediately brings the robot into primary focus and causes the robot to physically rotate and pan to bring the corresponding point into center focus. A vertical scanning Hokuyo provides depth estimates for objects in center focus, which are presented as candidates within this area. Samplings of these depth values are overlaid in the front camera image (Fig. 20d). For candidates outside the center focus, the patch height is used to estimate depth in meters by interpolating the height-depth relationship of the OOI. These depth estimates as well as the current angle of the front camera are used to determine the position of the OOI relative to the robot and ultimately its position in the global map. In addition to the image views, the screen contains control buttons (Fig. 20e), each of which has a corresponding key-binding, and indicators denoting which robot in primary focus will be commanded.

The other screen is less frequently used and contains views from non-selected robots (Fig. 20h), sliders (Fig. 20i) for manually adjusting camera parameters in case of extreme lighting conditions and two types of history images. To help compensate for the operator's limited attention and the quickly changing imagery of the auto-exploring robots, if the operator catches a flash of a potential candidate as it moves out of view, the past 5 seconds of imagery can be viewed for a selected robot, shown on top of the screen (Fig. 20f). Below this sequence, OOI history is presented (Fig. 20g). Each discovered OOI is assigned a unique ID, which can be used to initiate a neutralization or renounce a misidentification. These images are used to associate each OOI with their iconic information in the scene for reference.

The design of our system contrasts with other approaches in several ways. The first is the wealth of visual information streamed. Each robot in our team contained two dedicated cameras, with an omnidirectional image for the general surroundings and a front mounted panning one used to focus on visual object details. These images enabled the team to monitor for threats in all directions while simultaneously detecting and investigating OOIs in the periphery. The second is how this information is presented to the user. Rather than artificially limiting the information provided to the human operator, the visual information is organized in a manner that allowed for the operator to attend in a foveal manner to certain tasks while still maintaining a omniprescient view of the world. This highlights a key challenge in striking the proper balance between information overload and resource management in human-robot interfaces, especially as the number of robots and responsibilities of the human operator grows.

8.3 Ground control

The Strategy/Plan operator is responsible for monitoring and coordinating the strategic tasks of the individual UGVs in the robot team. For this purpose, we constructed an easy-to-use interface to control and modify the tasks of individual UGVs as well as groups of UGVs. Shown in Figure 21 is a user interface for the Strategy/Plan operator that is modeled after a real-time strategy (RTS) game. By pointing and clicking on various locations in the global map, the human operator can quickly retask and direct individual UGVs using very high-level commands to the underlying planners and

controllers in our system.

These tasks and behaviors use MATLAB data formats and scripts to relay human operator commands to the software modules on the different robots. Through the continuously updated real-time map display, the operator can quickly monitor the progress of the robot team and efficiently verify and possibly intercede during mission-critical operations.

The architecture of our overall system has hooks for the human operator to monitor the status of any incoming and outgoing messages between the various software modules. Human intervention is needed when there is large uncertainty in the information being passed between modules. For example, if there is uncertainty in the identity or location of an OOI, the OOI operator is notified to verify the putative identify and location before any further operation can continue. After the correct location is confirmed, the OOI is automatically displayed in the global map to the Strategy/Plan operator to verify whether a neutralization procedure can be initiated. With a single click on the user interface, the Strategy/Plan operator allows the system to autonomously task the nearest available UGVs to plan and coordinate their movements and actions to neutralize the identified OOI. Thus, the key design approach in our system can be characterized as being "certain about uncertainty" regarding any operational aspect that requires human intervention.

8.4 Mission operations

Our human-machine interface design allows the operators to quickly direct and guide operational policies by the UGV robot team. The ground control station can be used to define particularly relevant areas to explore, and any potentially dangerous areas to avoid. To determine these areas, we first identify areas of interest during a pre-operations stage where analysis of overhead imagery data is incorporated. These areas can then be preloaded into operational configuration files for use during particular phases of the mission.

Once the operational phase has begun, the global map and metadata feed overlay are utilized to quickly modify operational policies for the UGVs by the Strategy/Plan operator. As the OOI operator provides visual situational awareness of the surrounding environment, directed response

actions by the UGVs are coordinated by the two human operators.

9 Integration with UAVs

Our recent work has focused on incorporating aerial quadrotor robots to form truly hetergeneous robot teams capable of long term surveillance in difficult environments. Aerial robots have the ability to fly above terrains too difficult to traverse, and to obtain a full range of 3D perpective sensing data. Unfortunately, they have limited batteries and short flight times. We have merged the capabilities of aerial robots with ground vehicles by developing a magnetic landing mechanism as shown in Figure 22. Our experiments show how to process the information collected by such teams of robots.

10 Summary

This report has presented the technical approach of the University of Pennsylvania team during this project. We designed and constructed a large team of UGVs, along with recent work with UAVs, with an appropriate set of computational, sensing, communications and actuation hardware. We use a hierarchical series of sensing, planning, and control modules to coordinate and direct the autonomous navigation and actions of the robot team to achieve search, mapping, and neutralization mission objectives.

On each robot, a complementary set of sensor readings is fused to produce static and dynamic maps of the surrounding environment. By properly representing uncertainty and integrating the inertial, odometry, visual and LIDAR measurements, each individual UGV constructs a local map without GPS. These local maps are then hierarchically merged at the ground control station with metadata from overhead imagery and the simulated UAV feed to construct a globally consistent map of the robot team activities. This global map is displayed to the Strategy/Plan human operator who can then issue appropriate high-level strategic commands to the UGV team. Visual imagery from the robot omnidirectional and frontal view cameras are used by the second human operator to verify the identities and locations of OOIs, and provides the human operators with good situational

awareness during mission-critical events.

References

Bohren, J., Foote, T., Keller, J., Kushleyev, A., Lee, D. D., Stewart, A., Vernaza, P., Derenick, J. C., Spletzer, J. R., and Satterfield, B. (2008). Little Ben: The Ben Franklin racing team's entry in the 2007 DARPA Urban Challenge. *Journal of Field Robotics*, 25:598–614.

Butzke, J. and Likhachev, M. (2011). Planning for multi-robot exploration with multiple objective utility function. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3254–3259.

DSTO (2010). http://www.dsto.defence.gov.au/MAGIC2010/.

Edelman, S. (1997). Computational theories of object recognition. In *Trends in Cognitive Science*, pages 296–304.

Ferguson, D., Howard, T., and Likhachev, M. (2008). Motion planning in urban environments. *Journal of Field Robotics*, 25:939–960.

Geyer, C. and Daniilidis, K. (2003). Omnidirectional video. *The Visual Computer*, 19:405–416.

Henry, P., Krainin, M., Herbst, E., Ren, X., and Fox, D. (2012). RGB-D mapping: Using kinect-style depth cameras for dense 3d modeling of indoor environments. *International Journal of Robotics Research*, 31:647–663.

Hsieh, M. A., Cowley, A., Keller, J. F., Chaimowicz, L., Grocholsky, B., Kumar, V., Taylor, C. J., Endo, Y., Arkin, R. C., Jung, B., Wolf, D. F., Sukhatme, G. S., and MacKenzie, D. C. (2007). Adaptive teams of autonomous aerial and ground robots for situational awareness. *Journal of Field Robotics*, 24:991–1014.

Likhachev, M. and Ferguson, D. (2008). Planning long dynamically-feasible maneuvers for autonomous vehicles. In *Proceedings of Robotics: Science and Systems*. MIT Press.

Likhachev, M., Gordon, G., and Thrun, S. (2003). ARA*: Anytime A* with provable bounds on sub-optimality. In *Advances in Neural Information Processing Systems*, volume 16. MIT Press.

Michael, N., Fink, J., and Kumar, V. (2008). Experimental testbed for large multi-robot teams: verification and validation. *IEEE Robotics and Automation Magazine*, pages 53–61.

Naroditsky, O., Patterson, A., and Daniilidis, K. (2011). Automatic alignment of a camera with a line scan LIDAR system. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 3429–3434.

Olson, E. (2008). *Robust and Efficient Robotic Mapping*. PhD thesis, MIT, Cambridge, MA.

Reid, R. and Bräunl, T. (2011). Large-scale multi-robot mapping in MAGIC 2010. In *Robotics, Automation and Mechatronics*, pages 239–244.

Simmons, R. and Apfelbaum, D. (1998). A task description language for robot control. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*.

Simmons, R., Apfelbaum, D., Burgard, W., Fox, D., Moors, M., Thrun, S., and Younes, H. L. S. (2000). Coordination for multi-robot exploration and mapping. In *Proceedings of the Seventeenth National Conference on Artificial Intelligence and Twelfth Conference on Innovative Applications of Artificial Intelligence*, pages 852–858.

Trimpe, S. and DAndrea, R. (2010). Accelerometer-based tilt estimation of a rigid body with only rotational degrees of freedom. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 2630–2636.

Vanrullen, R., Reddy, L., and Koch, C. (2004). Visual search and dual tasks reveal two distinct attentional resources. *Journal of Cognitive Neuroscience*, 16:4–14.

Vernaza, P. and Lee, D. D. (2006). Rao-Blackwellized particle filtering for 6-DOF estimation of attitude and position via GPS and inertial sensors. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 1571–1578.

Vernaza, P., Taskar, B., and Lee, D. D. (2008). Online, self-supervised terrain classification via discriminatively trained submodular Markov random fields. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 2750–2757.

Viola, P. and Jones, M. (2001). Robust real-time object detection. In *Proceedings of the International Conference on Computer Vision*, page 747.

Visser, A. and Slamet, B. (2008). Balancing the information gain against the movement cost for multi-robot frontier exploration. In *European Robotics Symposium*, volume 44, pages 43–52.

Yamauchi, B. (1997). A frontier-based approach for autonomous exploration. In *Proceedings of the Computational Intelligence in Robotics and Automation*.

Zlot, R., Stentz, A., Dias, M., and Thayer, S. (2002). Multi-robot exploration controlled by a market economy. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 3016–3023.

Figure 6: Detailed view of the sensor suite consisting of a panning high resolution camera, omni-directional camera, horizontal and vertical LIDAR scanners.

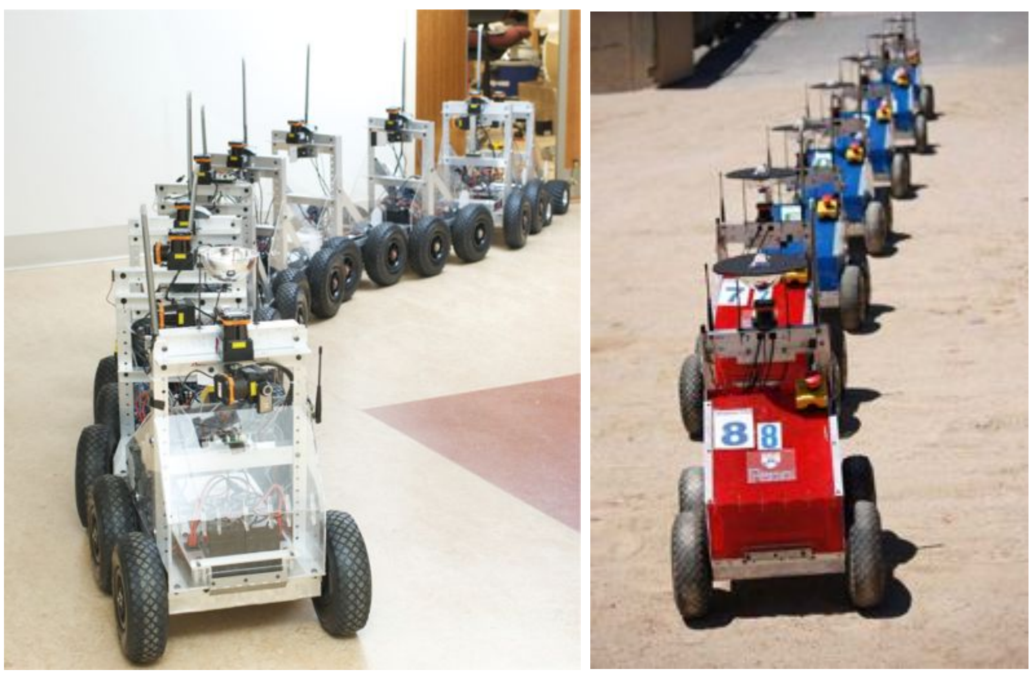
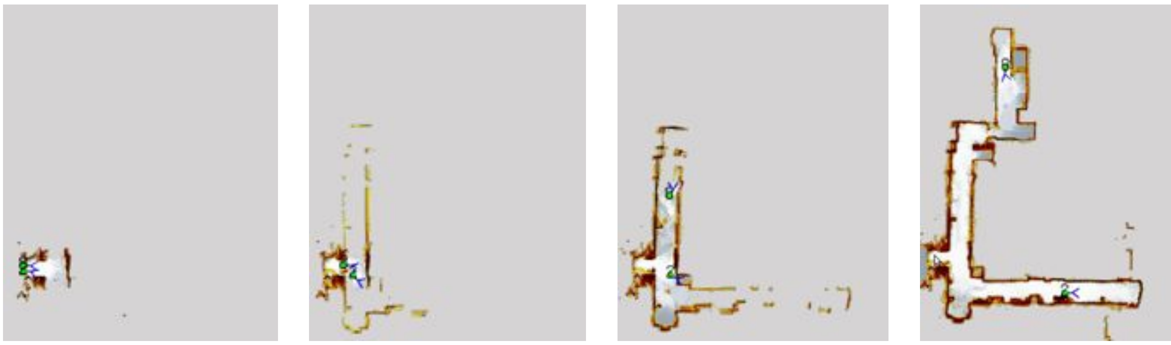


Figure 7: Nine UGVs were constructed and fielded in the competition. Two operators efficiently monitored and coordinated the team using a custom high-level graphical interface.



(a) t = 0s (b) t = 21s (c) t = 39s (d) t = 245s

Figure 8: An example of indoor exploration with two ground robots. Gray represents unknown area; dark colors are used for obstacles; explored ground is white.

(a) single robot exploration (b) exploration with five robots

Figure 9: Frontiers in an exploration task with one robot (a). Black represents obstacles; green is traversable ground; gray is unknown; borders between gray and green colors represents the frontiers. Exploration map with five robots is shown in (b). Exploration frontiers are highlighted with orange ellipses.

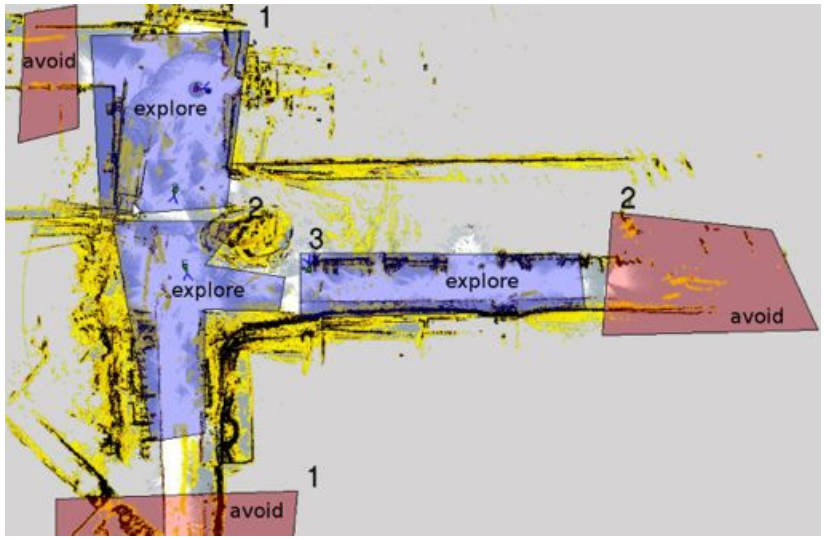


Figure 10: High-level user input allows to easily coordinate a group of robots on an exploration mission. User-selected avoid zones shown in red, exploration zones shown in blue.

Figure 11: Static OOI is found, its size and position with respect to the robot is calculated and displayed (2.0 m to the right and 11.9 m forward).

(a) mobile OOI (b) neutralization by tracking with two robots

Figure 12: Dynamic OOI neutralization. OOI apears in a red coveralls (a) and two robots track it with the cameras, simultaneously transmitting image data to the base station (b).

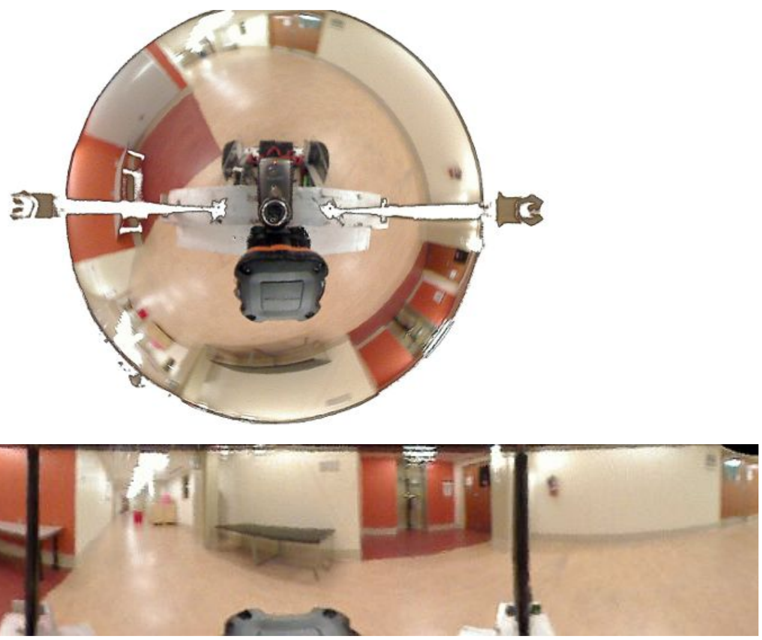
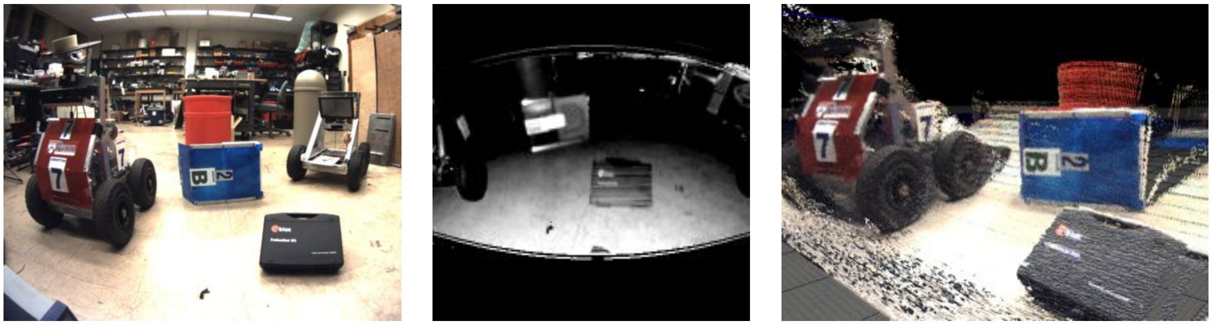
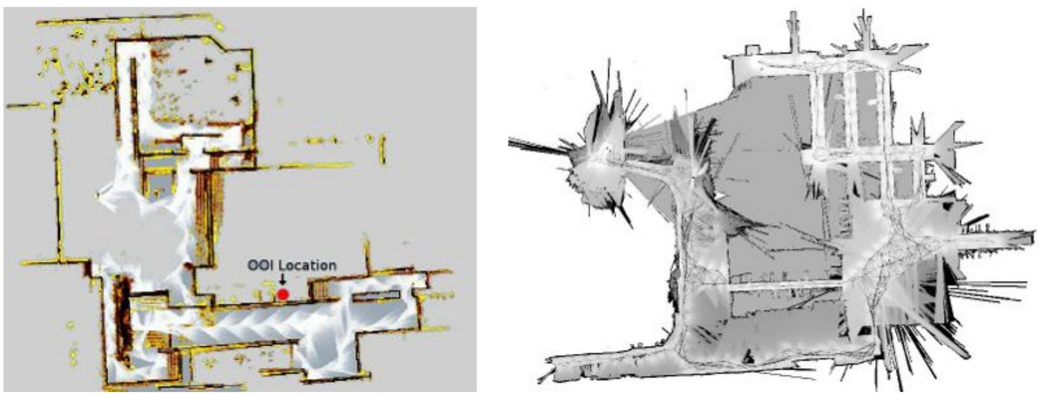


Figure 13: Omnidirectional image unwarped to provide 360 degree field of view around each UGV.



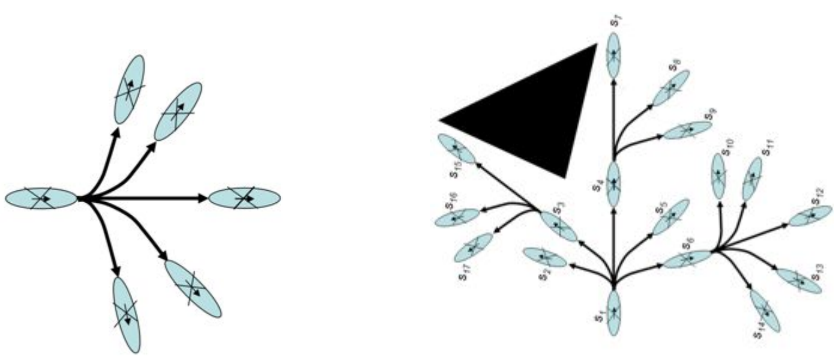
(a) raw camera image (b) LIDAR depth and intensity (c) resulting colored 3D point cloud

Figure 14: LIDAR-camera calibration of extrinsic sensor parameters.



(a) Indoor map built by a single robot (b) Outdoor map built by a team of seven robots

Figure 15: Two complementary LIDAR sensors are used to map both traversable ground and obstacles simultaneously. Dark colors represent unknown terrain and obstacles; light colors represent traversable ground.



(a) an example of a motion template (b) graph constructed using motion templates

Figure 16: Sample action template (a) and part of the resulting lattice graph (b). Black triangle represents an obstacle and actions that result in collision with the obstacle receive high cost and are effectively pruned.

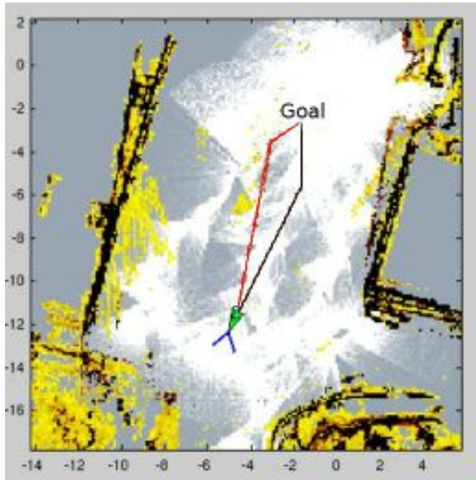


Figure 17: Global (red) and local (black) plans. Global plan is coarse and goes through some uncertain areas. Local plan is much better and tells the robot to go through known and traversable terrain (white).

(a) robot entering maze environment

(b) final map built with five robots

Figure 18: Indoor mapping of the Old Ram Shed Challenge environments using five robots. Robot tracks in (b) are shown as thin dark lines.

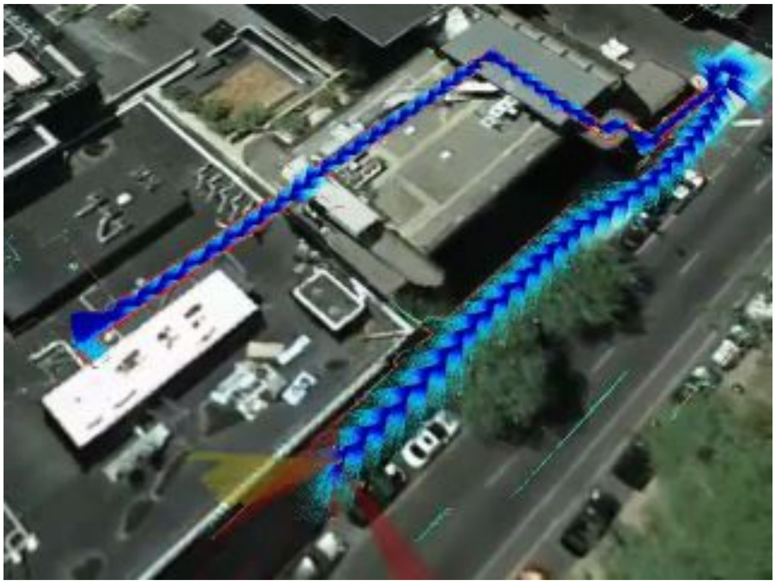


Figure 19: Registering UGV indoor and outdoor map information with overhead imagery of the UPenn campus environment.

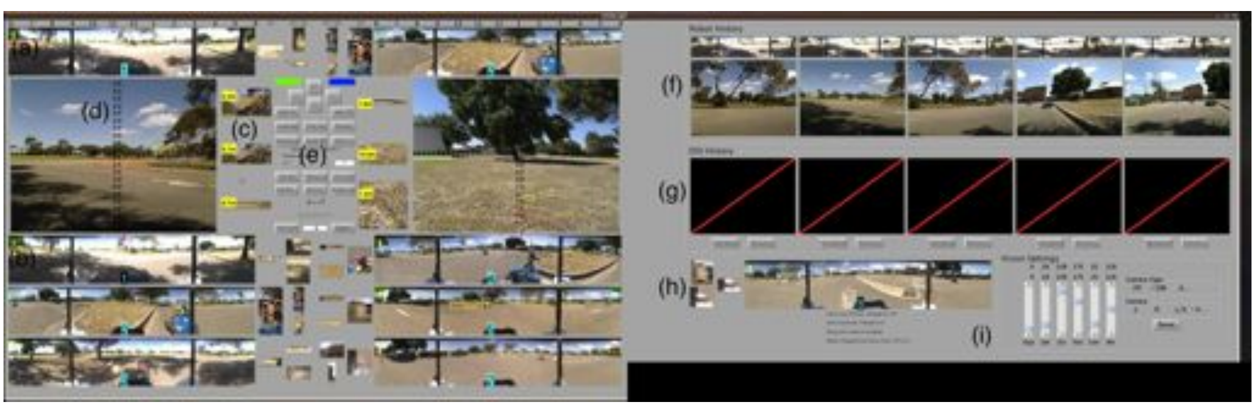


Figure 20: (a) Two robots being monitored with omni and front view images (green); (b) omni views of robots not in focus (blue); (c) enlarged OOI candidate patches (yellow); (d) image mapped depth values; (e) control buttons; (f) front view history of selected robot; (g) OOI history and scene context; (h) omni view of last used robot; (i) camera parameter controls.

Figure 21: Ground control display of the Strategy/Plan operator with intuitive human-machine interface controls.

Figure 22: Magnetic tailhook system used to assist aerial quadrotor landing and mating on a ground robot.